



Glottal opening and closing events investigated by electroglottography and super-high-speed video recordings

Christian T Herbst, Jörg Lohscheller, Jan G Švec, Nathalie Henrich
Bernardoni, Gerald Weissengruber, W Tecumseh Fitch

► To cite this version:

Christian T Herbst, Jörg Lohscheller, Jan G Švec, Nathalie Henrich Bernardoni, Gerald Weissengruber, et al.. Glottal opening and closing events investigated by electroglottography and super-high-speed video recordings. *Journal of Experimental Biology*, 2014, 217 (Pt 6), pp.955-963. 10.1242/jeb.093203 . hal-01015251

HAL Id: hal-01015251

<https://hal.science/hal-01015251>

Submitted on 26 Jun 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RESEARCH ARTICLE

Glottal opening and closing events investigated by electroglottography and super-high-speed video recordings

Christian T. Herbst^{1,2,*}, Jörg Lohscheller³, Jan G. Švec¹, Nathalie Henrich⁴, Gerald Weissengruber⁵ and W. Tecumseh Fitch²

ABSTRACT

Previous research has suggested that the peaks in the first derivative (dEGG) of the electroglottographic (EGG) signal are good approximate indicators of the events of glottal opening and closing. These findings were based on high-speed video (HSV) recordings with frame rates 10 times lower than the sampling frequencies of the corresponding EGG data. The present study attempts to corroborate these previous findings, utilizing super-HSV recordings. The HSV and EGG recordings (sampled at 27 and 44 kHz, respectively) of an excised canine larynx phonation were synchronized by an external TTL signal to within 0.037 ms. Data were analyzed by means of glottovibrograms, digital kymograms, the glottal area waveform and the vocal fold contact length (VFCL), a new parameter representing the time-varying degree of 'zippering' closure along the anterior–posterior (A–P) glottal axis. The temporal offsets between glottal events (depicted in the HSV recordings) and dEGG peaks in the opening and closing phase of glottal vibration ranged from 0.02 to 0.61 ms, amounting to 0.24–10.88% of the respective glottal cycle durations. All dEGG double peaks coincided with vibratory A–P phase differences. In two out of the three analyzed video sequences, peaks in the first derivative of the VFCL coincided with dEGG peaks, again co-occurring with A–P phase differences. The findings suggest that dEGG peaks do not always coincide with the events of glottal closure and initial opening. Vocal fold contacting and de-contacting do not occur at infinitesimally small instants of time, but extend over a certain interval, particularly under the influence of A–P phase differences.

KEY WORDS: Electroglottography, dEGG double peaks, High-speed video, Vocal fold contact length, Excised canine larynx, Glottal opening and closing events

INTRODUCTION

Monitoring and assessment of vocal fold vibration in voice production is crucial for understanding normal and disordered voice, treating voice disorders and training professional voice users. Observation of the larynx during phonation is performed with either direct or indirect methods. The direct methods, such as videostrobolaryngoscopy (Bless et al., 1987), videokymography

(Švec and Schutte, 1996) or high-speed videoendoscopy (Rubin and LeCover, 1960; Moore et al., 1962; Hertegård, 2005; Deliyski et al., 2008; Deliyski and Hillman, 2010), provide insights into the spatio-temporal oscillatory behaviour of laryngeal tissue. However, they are semi-invasive (not well tolerated by some subjects), cost intensive, and usually only performed by trained personnel in dedicated premises such as voice clinics.

As a low-cost, non-invasive alternative, vocal fold vibration can be monitored indirectly by electroglottography (EGG) (Fabre, 1957). A low-amperage, high-frequency current is passed between two electrodes placed on each side of the thyroid cartilage at vocal fold level. The time-varying change of vocal fold contact during the flow-induced oscillation of laryngeal tissue induces variations in the electrical impedance across the larynx, resulting in variation in the current between the two electrodes (Fourcin and Abberton, 1971; Baken, 1992; Baken and Orlikoff, 2000). These admittance variations are proportional to the relative vocal fold contact area during phonation (Scherer et al., 1988).

Experimental research has suggested that landmarks in the EGG waveform are related to the relative movement and position of the vocal folds during phonation (Rothenberg, 1979; Baer et al., 1983; Childers et al., 1983; Hess and Ludwigs, 2000). The physiologic relevance of the EGG signal has been examined theoretically by Childers et al. (Childers et al., 1986) and Titze (Titze, 1989; Titze, 1990), the latter discussing the effects of (1) increased glottal adduction, (2) glottal convergence (with vertical phasing), (3) medial vocal fold surface bulging and (4) increased vertical phasing in vocal fold vibration.

The moments of glottal opening and glottal closure are of particular interest for quantitative analysis of the voice source. The timing of these events can be used to determine the relative proportion of glottal closure within a glottal vibratory period (Rothenberg and Mahshie, 1988), known as the 'larynx closed quotient' (Howard, 1995) or 'contact quotient' [CQ_{EGG} (Orlikoff, 1991)]. This quotient has been found useful in clinical as well as in basic voice research (e.g. Schutte and Miller, 2001; Henrich et al., 2005; Švec et al., 2008). However, the calculation of the CQ_{EGG} is influenced by the choice of algorithm used to determine the contacting and de-contacting instants, and must therefore be used with caution (Sapienza et al., 1998; Higgins and Schulte, 2002; Henrich et al., 2004; Kania et al., 2004; Herbst and Ternström, 2006; La and Sundberg, 2012).

For the purpose of calculating the CQ_{EGG}, estimation of the instants of glottal closure and opening is performed by either (1) applying a threshold criterion to the locally normalized EGG signal (Rothenberg and Mahshie, 1988), or (2) finding positive and negative maxima in the first mathematical derivative of the EGG signal (dEGG), reflecting the maximum rate of change of the EGG signal with time (Teaney and Fourcin, 1980; Childers and Krishnamurthy, 1985; Henrich et al., 2004). Because the latter

¹Voice Research Laboratory, Department of Biophysics, Faculty of Science, Palacký University Olomouc, tr. 17. Listopadu 12, 771 46 Olomouc, Czech Republic. ²Laboratory of Bio-Acoustics, Department of Cognitive Biology, University of Vienna, Althanstraße 14, 1090 Vienna, Austria. ³University of Applied Sciences, Department of Computer Science, Schneidershof, 54293 Trier, Germany. ⁴GIPSA-lab, CNRS, Grenoble INP, Grenoble University, 11 rue des Mathématiques – BP 46, 38402 Saint Martin d'Hères cedex, France. ⁵University of Veterinary Medicine Vienna, Institute for Anatomy, Histology and Embryology, Veterinärplatz 1, 1210 Vienna, Austria.

*Author for correspondence (christian.herbst@univie.ac.at)

List of abbreviations	
A–P	anterior–posterior
CQ _{EGG}	electroglottographic contact quotient
dEGG	first derivative of the electroglottographic signal
dGAW	first derivative of the glottal area waveform
dVFCL	first derivative of the vocal fold contact length signal
DKG	digital kymography or digital kymogram
EGG	electroglottography or electroglottographic
GAW	glottal area waveform
GVG	glottovibrogram
HSV	high-speed video
MCQ	membranous contact quotient
TTL	transistor-transistor logic
VFCL	vocal fold contact length

approach does not rely on arbitrary user input (i.e. the arbitrary choice of a threshold) but is rather based on intrinsic properties of the EGG signal, it may be better suited for detection of the glottal opening and closure instants. The relation of positive and negative peaks in the dEGG signal to the events of glottal closure and opening, as seen in laryngeal imaging performed with a wide range of imaging frame rates, has been evaluated in several studies (see Table 1). These studies highlight the potential of the one-dimensional EGG signal (and in particular its first derivative, dEGG) to reveal information about the complex three-dimensional contacting motion of the vocal folds in a non-invasive fashion.

The vocal folds do not vibrate as a uniform mass [e.g. as is seen in a one-mass model of the vocal folds (see Flanagan and Landgraf, 1968)]. Rather, their vibration is characterized by phase differences along both the inferior–superior (Baer, 1981; Titze et al., 1993) and anterior–posterior (A–P) dimensions (Tanabe et al., 1975; Krenmayr et al., 2012; Orlikoff et al., 2012; Yamauchi et al., 2013). These phase differences cause time-delayed contacting and de-contacting of the vocal folds along the respective axes. There is thus no specific instant of glottal closing and opening, but rather an interval during which the closing and opening, respectively, occur. This is reflected in some of the findings summarized in Table 1. Furthermore, phase differences along the A–P dimension, i.e. the so-called ‘zipper-like’ opening and closure (Childers et al., 1986), may introduce multiple peaks into the dEGG waveform (Hess and Ludwigs, 2000; Henrich et al., 2004). The data reported in a recent study (Herbst et al., 2010) give reason to assume that multiple dEGG peaks represent systematic physiological phenomena rather than artifacts.

In most of the studies shown in Table 1, the frame rate of the high-speed video (HSV) recordings was approximately an order of magnitude below the commonly used sampling frequency for recording EGG data (i.e. 44,100 Hz). The timing accuracy of assessments of glottal closing and opening instants or intervals is thus limited by the lower video frame rate. In particular, what appears to be a closing instant in HSV with a low frame rate might actually turn out to be a closing interval in HSV with a higher video frame rate, particularly if the glottal opening or closing exhibits a phase delay along the A–P glottal axis. In order to investigate this issue, super-HSV imaging with a video frame rate of more than half the EGG signal sampling rate is used here to relate the landmarks of vocal fold vibration to those found in the dEGG signal.

RESULTS

The process used to extract the three analyzed sequences and their respective subglottal pressures is illustrated in Fig. 1.

Sequence 1: double dEGG peaks in de-contacting phase

The electroglottographic (EGG) and vibratory data for sequence 1 are illustrated in Fig. 2. In the opening phase, a pronounced A–P phase difference (‘zippering’) was seen following full glottal closure from ~5.2 ms to ~8.0 ms (see arrow in Fig. 2D, supplementary material Movie 1), suggesting the presence of an x-20 or x-21 vocal fold vibratory mode (Berry et al., 1994) (see Fig. 3 and supplementary material Movie 8). The EGG signal amplitude started to decrease ~2 ms before the moment of initial glottal opening (marker T1), which suggests the presence of a strong phase difference of the vocal fold vibration along the inferior–superior dimension. The inferior vocal fold edges – not seen in the HSV – presumably started to separate around $t \approx 3$ ms, just after the previous moment of complete glottal closure.

Two negative maxima of similar amplitude were found in the dEGG (see dashed vertical markers T1 and T2 in Fig. 2A). The first of these negative peaks was synchronized with the moment of initial glottal opening, preceding it by only 0.02 ms (see supplementary material Movie 2). This event was also reflected by a local maximum in the derivative of the glottal area waveform (dGAW) and a pronounced local minimum in the derivative of the vocal fold contact length waveform (dVFCL; see Fig. 2B,C). The second negative dEGG peak (marker T2) occurred when the posterior glottis was still partially closed, and it did not coincide with a peak in either the dGAW or dVFCL waveform.

Table 1. Overview of six studies relating positive and negative peaks in the dEGG signal to glottal closing and opening events as seen in laryngeal imaging

Study	Video frame rate (frames s ⁻¹)	Finding
Baer et al., 1983	4000	Sharp dEGG peak in contacting phase, which ‘occurs over an interval that is comparable to the interval between film frames’; glottal opening may not be a single event: it ‘occurs at different times along the anterior–posterior extent of the glottis’
Childers et al., 1983	5000	Glottal closure may be predicted from dEGG; instant of glottal opening is more variable, but seems to occur the instant at which the dEGG waveform has a maximum; mucus bridges may cause artefacts
Berke et al., 1987	30 (videostroboscopy)	Contacting dEGG peak in Berke’s data corresponded to complete closure of the lower vocal fold margin; no landmark corresponded to decontacting dEGG peak
Anastaplo and Karnell, 1988	30 (videostroboscopy)	‘Opening along the superior margin of the glottis usually coincides with the knee observed during opening phase of the EGG waveform’
Henrich et al., 2004	3704	‘dEGG peaks may be considered as reliable indicators of glottal opening and closing instants’
Orlikoff et al., 2012	16,000	In previous studies, ‘peaks in the dEGG have been shown to correspond well to glottal opening and closing instants’; some EGG slope changes may be related to anterior–posterior phase differences (as opposed to superior-inferior phase differences)

EGG, electroglottographic; dEGG, the first derivative of the EGG signal.

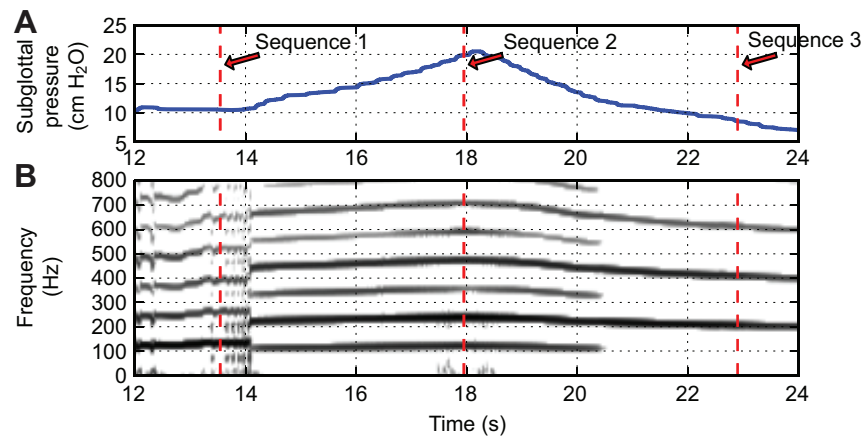


Fig. 1. Overview of sequence extraction from the excised larynx pressure sweep. (A) Trace of sub-glottal pressure during pressure sweep. A total of three sequences (each representing two complete periods of vocal fold vibration) were extracted from locally stable regions having a minimum of 20 similar periods of oscillation (in the case of period doubling, two consecutive phases of vocal fold contacting and de-contacting were counted as one period), based on analysis of the electroglottographic (EGG) signal and the first derivative of the EGG signal (dEGG). Sequence 1, pronounced double peaks in the de-contacting phase (≈ 13.59 s, duration ≈ 15.6 ms); sequence 2, clear single peaks in both the contacting and the de-contacting phase (≈ 17.95 s, duration ≈ 17.0 ms); sequence 3, pronounced double peaks in the contacting phase (≈ 22.90 s, duration ≈ 9.9 ms). (B) Narrow-band spectrogram of the EGG signal, window duration ≈ 93 ms.

The closing phase was characterized by an A–P phase difference (‘zippering’), suggesting again that the x-20 or x-21 mode contributed to the vibratory pattern. One distinct dEGG maximum was observed (Fig. 2A, marker T3), which preceded the moment of complete glottal closure (HSV, marker T4) by 0.61 ms.

The kymograms in Fig. 2E–H depict the time-varying glottal opening at 80, 60, 40 and 20% of the entire glottal length, respectively. The quantitative glottal width data, extracted from the glottovibrogram (GVG) data in Fig. 2D, were superimposed upon one complete glottal cycle (see the light blue shapes in Fig. 2E–H). The A–P phase difference in both the opening and closing phases was reflected in the kymograms. Marker T1 coincided with the moment of initial glottal opening at a position of 20% of the glottal length, i.e. digital kymogram (DKG) 0.2 in Fig. 2H. With increasing posterior position along the glottal axis, the duration of the glottal open phase decreased. Consequently, the DKG 0.8 (Fig. 2E) location, extracted at a position closest to the posterior boundary of the glottis, had the longest closure duration.

Sequence 2: single dEGG peaks

The EGG and vibratory data for sequence 2 are shown in Fig. 4. Because of the observed period doubling in this sequence, each period contained two glottal cycles, and each of these cycles consisted of one phase of vocal fold de-contacting and contacting, respectively. These two cycles are identified in Fig. 4 as ‘cycle 1’ (from marker T0 to marker T3) and ‘cycle 2’ (from marker T3 to marker T6).

The opening phase of cycle 1 was characterized by a slight GVG ‘hourglass’ pattern (see arrows in Fig. 4D, supplementary material Movie 3) occurring over a period of ~ 0.2 ms, suggesting the presence of an x-30 or x-31 vibratory mode (see Fig. 3, supplementary material Movie 9). The decrease of the EGG signal amplitude occurred over a duration of ~ 1 ms just before the moment of initial glottal opening (Fig. 4A, vertical marker T1). As in the previous sequence, this may indicate the presence of a phase difference of the vocal fold vibration along the inferior–superior dimension. The inferior vocal fold edges – not seen in the HSV – presumably started to separate around ≈ 3 ms (or even slightly earlier, if the decrease of vocal fold contact area at the inferior vocal fold margin was counteracted by an increase of vocal fold contact

area along the superior vocal fold margin, resulting in a ‘flat-top’ EGG waveform between ≈ 2 ms and ≈ 3 ms). One distinct negative dEGG peak (Fig. 4A) was found in the de-contacting phase (dashed vertical marker T2 in Fig. 4), which was delayed by 0.13 ms from the moment of initial glottal opening in the HSV data (see supplementary material Movie 4). This dEGG peak was temporally aligned with a local maximum of the dGAW waveform (Fig. 4B) and a local minimum of the dVFCL waveform (Fig. 4C). These peaks occurred at the moment when the central portion of the glottis opened, i.e. when the vocal fold edges lost their contact along the entire glottal axis (marker T2).

In cycle 1, the glottis closed with a slight ‘anti-hourglass’ zippering motion towards the center of the glottal axis (see supplementary material Movie 3), again suggesting the presence of an x-31 vibratory mode. One pronounced positive dEGG peak was found (Fig. 4A, vertical marker T3). This peak was temporally aligned with the moment of complete glottal closure (as determined from the GVG, Fig. 4D) and a positive peak in the dVFCL waveform (Fig. 4C).

The opening phase of cycle 2 was also characterized by a slight ‘hourglass’ pattern (see supplementary material Movie 3), suggesting the presence of an x-30 or x-31 vibratory mode. The moment of initial glottal opening was reflected by a negative peak in the dVFCL waveform (Fig. 4C, vertical marker T4). One pronounced negative peak was found in the dEGG signal (Fig. 4A, vertical marker T5), which was delayed by 0.45 ms as compared with the moment of initial glottal opening (marker T4). The negative dEGG peak coincided with a positive peak in the dGAW waveform and a negative peak in the dVFCL waveform (Fig. 4B,C, marker T5).

In cycle 2, the vocal folds closed with an ‘anti-hourglass’ zippering motion towards the center of the A–P glottal axis (see supplementary material Movie 3), again suggesting the continuous presence of an x-30 or x-31 vibratory mode. One pronounced positive peak was found in the dEGG waveform (Fig. 4A), which was synchronized with a positive peak of the dVFCL waveform (Fig. 4C) and preceded the moment of glottal closure as determined from the GVG (Fig. 4D, marker T6) by 0.12 ms.

The kymograms shown in Fig. 4E–H were derived from the HSV and GVG data in a similar fashion as those in Fig. 2. For cycle 1, the moments of initial glottal opening and complete glottal closure

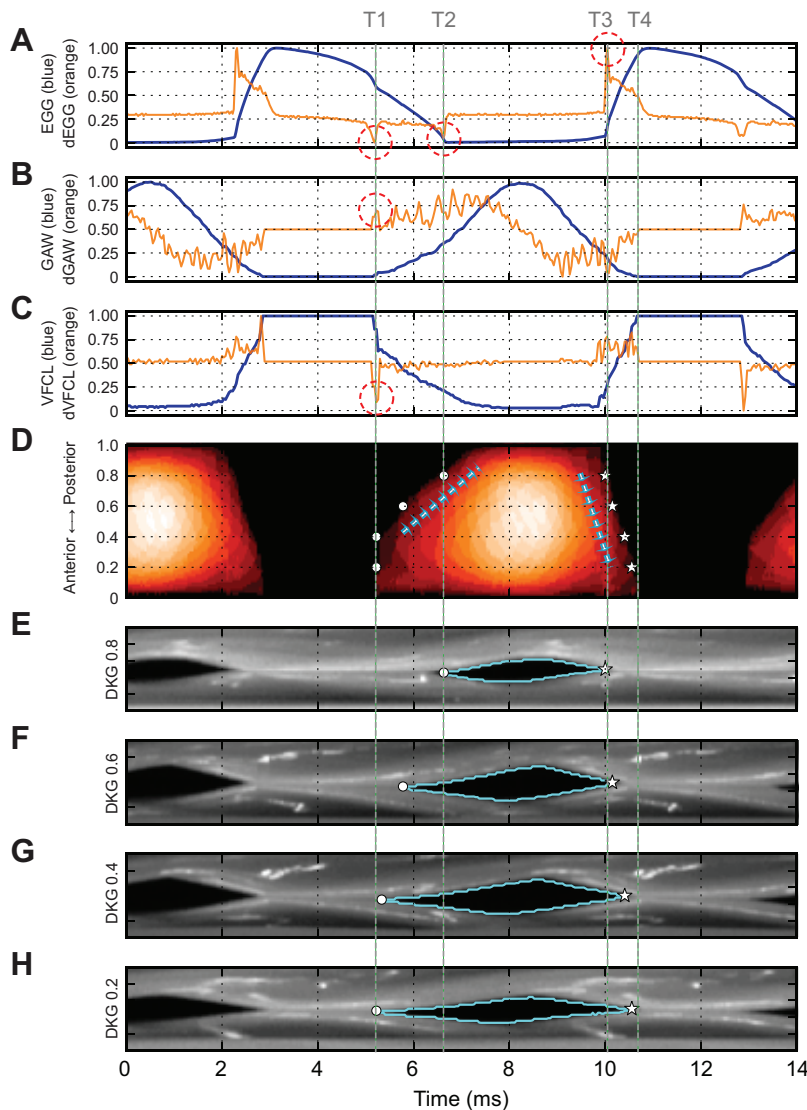


Fig. 2. Time-synchronous electroglottographic data and vibratory characteristics of sequence 1. (A–C) Locally normalized waveforms of EGG signal, glottal area waveform (GAW) and vocal fold contact length (VFCL) (blue), and the respective normalized first derivatives thereof: dEGG, dGAW and dVFCL (orange). The red dashed circles indicate intersections between peaks in the derivatives and the dashed vertical markers T1–T4. (D) Glottovibrogram (GVG), based on glottal edge extraction from high-speed video data (see Materials and methods). The light blue arrows depict the gradual glottal opening and closing, respectively, along the anterior–posterior glottal axis. (E–H) Digital kymographs (DKGs), extracted from high-speed video data at different positions along the anterior–posterior glottal axis. The extracted time-varying glottal edges are superimposed in light blue over one glottal cycle. The circles and stars in both the GVG and the DKGs indicate moments of glottal opening and glottal closure, respectively, as seen at different points along the anterior–posterior glottal axis. The gray dashed vertical markers T1–T4 highlight landmarks of the glottal cycle: T1, first negative dEGG peak, moment of initial glottal opening; T2, second negative dEGG peak; T3, positive dEGG peak; T4, moment of complete glottal closure.

(markers T1 and T3) were reflected in the DKG 0.2 (Fig. 4H) and DKG 0.6 (Fig. 4F), respectively. The moment of initial glottal opening in cycle 2 (marker T4) coincided with that in the DKG 0.2, and the moment of complete glottal closure in that cycle (marker T6) was reflected by DKG 0.4 (Fig. 4G).

Sequence 3: double dEGG peaks in contacting phase

The EGG and vibratory analysis data for sequence 3 are displayed in Fig. 5 (see also supplementary material Movie 5). The vocal fold vibration was characterized by a short closed phase of ~12% of the glottal cycle duration, as determined from the VFCL (Fig. 5C) and GVG data (Fig. 5D). The initial glottal opening occurred in the

anterior portion of the glottal axis, preceding the initial opening of the posterior glottis by more than 0.5 ms (see markers T1 and T3 in Fig. 5D). One strong minimum was found in the dEGG signal (Fig. 5A, dashed vertical marker T2), which lagged the moment of initial glottal opening (marker T1) by 0.49 ms. This negative dEGG peak, which occurred when the central portion of the glottis opened, did not coincide with any landmark in either the dGAW or dVFCL signals. The dGAW and dVFCL signals had one synchronized peak (Fig. 5B,C, marker T3) at the moment when the posterior portions of the vocal folds separated.

The closing phase of this sequence was characterized by an ‘anti-hourglass’ zipping (see supplementary material Movie 5)

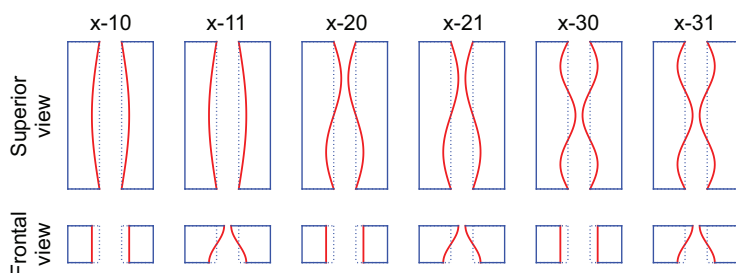


Fig. 3. Theoretical mode shapes of vocal fold vibration eigenmodes. In the x-ij notation, ‘x’ indicates oscillations along the lateral–medial direction, and the *ij* indices denote the number of oscillatory half-wavelengths occurring along the horizontal (inferior–superior) and vertical (anterior–posterior) dimensions of the vocal folds, respectively (adapted from Berry et al., 1994; Švec, 2000). An animated version of this figure can be found online at www.christian-herbst.org/media/. The theoretical contribution of the x-2n and x-3n modes to vocal fold vibration is demonstrated in the animations provided as supplementary material Movies 7–9.

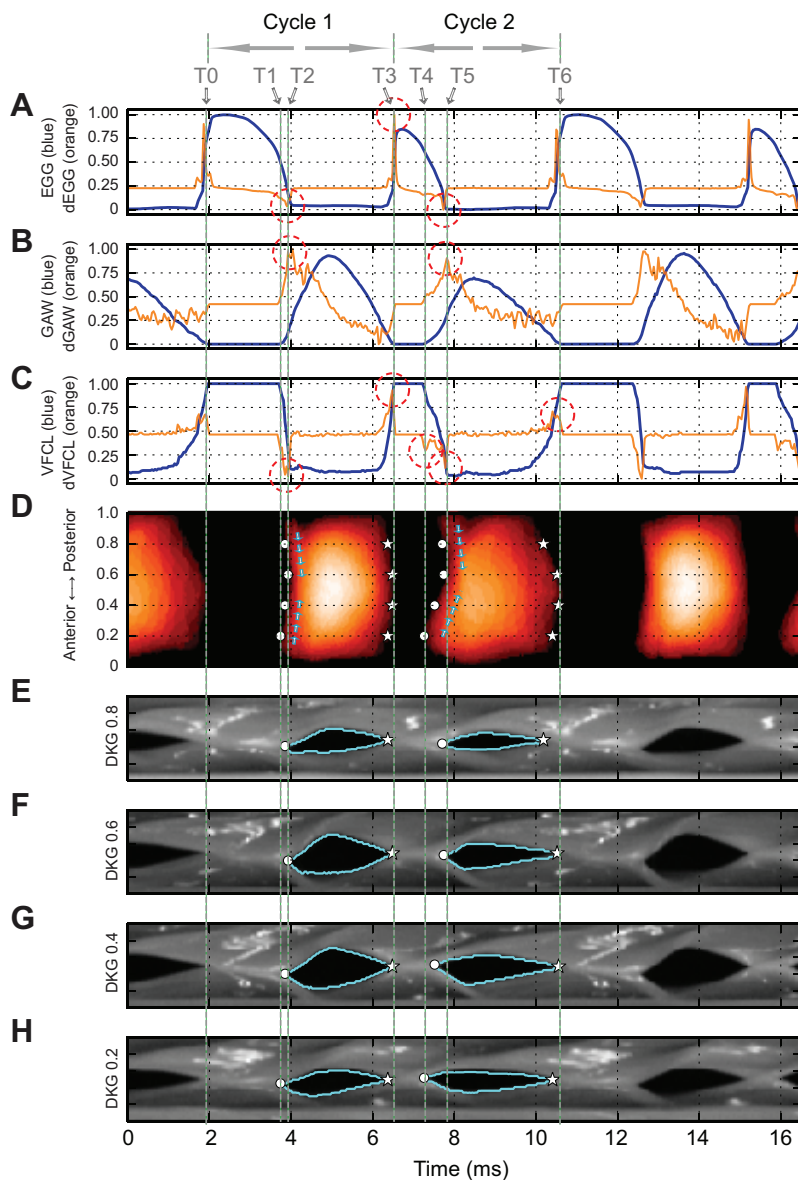


Fig. 4. Time-synchronous electroglottographic data and vibratory characteristics of sequence 2. See Fig. 2 for descriptions of all panels. The gray dashed vertical markers T0–T6 highlight landmarks of the glottal cycle: T0, beginning of cycle 1; T1 and T4, moments of initial glottal opening in cycles 1 and 2, respectively; T2 and T5, moments of complete separation of vocal folds in cycles 1 and 2, respectively; T3 and T6, moments of complete glottal closure in cycles 1 and 2, respectively.

towards the center of the A–P glottal axis, suggesting that also in this example an x-30 or x-31 vibratory mode participated in the vocal fold vibration (recall supplementary material Movie 9). Two distinct positive maxima were found in the dEGG signal (Fig. 5, markers T4 and T6), neither of which coincided with any other glottal landmark (see supplementary material Movie 6). The rise in the EGG waveform at marker T4 appears to involve a marked increase in tissue contact in the vertical plane not shown by the gradual GAW decrease (Fig. 5B) and VFCL increase (Fig. 5C) at that time.

The kymograms in Fig. 5E–H were derived from the HSV and GVG data in a similar fashion as those in Fig. 2. Because the moment of initial glottal opening (marker T1) occurred at a position of ca. 30% of the entire glottal length, it is not reflected in any of the displayed kymograms. The moment of complete glottal closure (marker T6) occurred at an offset of 40% along the glottal axis and was thus indicated in the DKG 0.4 (Fig. 5G).

The temporal offsets between glottal closing/opening events (as determined from HSV data) and the respective dEGG peaks for all analyzed sequences are summarized in Table 2.

DISCUSSION

In this study, we present high-speed data of vocal fold vibration recorded at a video frame rate of 27,000 frames s^{-1} . This is, to the best of our knowledge, the highest video frame rate reported in the literature to date for glottal observations, being almost seven times larger than the commonly available recordings made at rates of 4000 frames s^{-1} . The increased temporal accuracy was needed to gain better insights into the temporal alignment between glottal closing and opening events (as determined from HSV data) and positive and negative peaks found in the dEGG signal.

The hypothesis that dEGG waveform peaks provide clear indicators of glottal closing and opening instants was not supported in the three reported vibratory conditions. In only two out of eight cases was a good temporal agreement between dEGG peak and glottal event found: in the opening phase of sequence 1 (see Fig. 2) and the closing phase of cycle 1 in sequence 2 (see Fig. 4), with a measured delay of 0.02 ms, which was below the maximum synchronization error. In three cases, the closing or opening event, respectively, occurred ~ 0.5 ms before or after the occurrence of the respective dEGG peak (i.e. at an offset of 7–10% of the glottal cycle

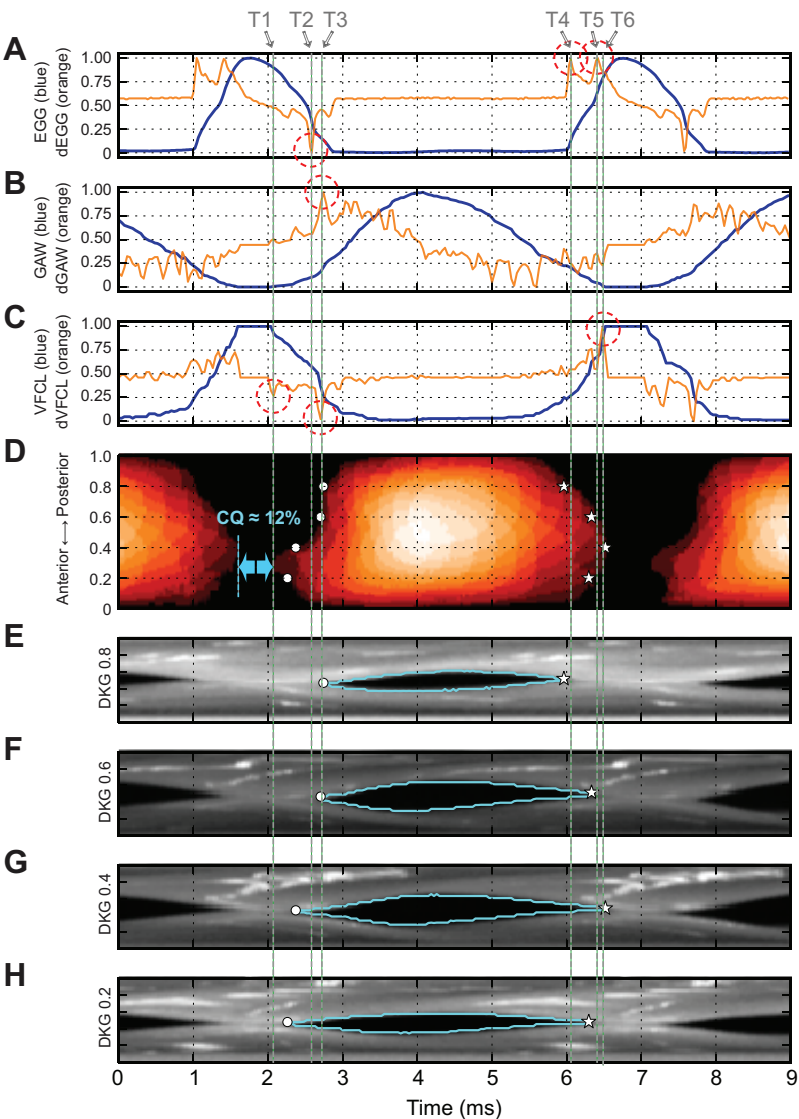


Fig. 5. Time-synchronous electroglottographic data and vibratory characteristics of sequence 3. See Fig. 2 for descriptions of all panels. The gray dashed vertical markers T1–T6 highlight landmarks of the glottal cycle: T1, moment of initial glottal opening; T2, negative dEGG peak; T3, peak in dGAW and dVFCL waveforms, coinciding with a simultaneous vocal fold separation along a third of the glottal axis; T4 and T5, positive dEGG peaks; T6, moment of complete glottal closure, coinciding with a positive dVFCL peak. CQ, closed quotient.

duration; see Table 2), suggesting that the dEGG peak did not coincide with the moment of glottal closing or opening, respectively. In the remaining three cases, the offset between dEGG peak and the respective glottal event was in the range of 0.1 to 0.15 ms. In sampled data, the smallest observable instant of time is determined by the sampling frequency (and consequently by the achievable exposure time in the case of video recordings) at which the analyzed data were acquired. As a rule of thumb, a maximum timing error of plus/minus half the time difference between two consecutive samples (i.e. half the synchronization error of 0.037 ms

in this study) should always be considered for estimating the accuracy of the occurrence of an instant in time, assuming that the observed vibratory phenomenon is not band limited [i.e. if frequency components higher than half the sampling frequency are present (see McClellan et al., 1998; Roads and Strawn, 1998)]. The data presented in this study show that increased video frame rates provide a means to better understand the relationship between EGG and HSV data (Golla et al., 2009), suggesting that it is not as simple as previously thought: vocal fold vibration may include phase delays along both the inferior–superior and the A–P glottal

Table 2. Overview of opening and closing events [as determined from high-speed video (HSV) data] in relation to the temporal position of the strongest dEGG peak in the corresponding de-contacting and contacting phases, respectively, for all three analyzed sequences

Sequence	Figure, glottal event markers	Cycle duration (ms)	Delay opening phase (ms) (% of cycle duration)	Delay closing phase (ms) (% of cycle duration)
1	Fig. 2, T1/T4	7.74	0.02 (0.24)	0.61 (7.88)
2 (cycle 1)	Fig. 4, T1/T3	4.56	–0.13 (–2.94)	–0.02 (–0.51)
2 (cycle 2)	Fig. 4, T4/T6	4.11	–0.45 (–10.88)	0.12 (2.91)
3	Fig. 5, T1/T6	4.93	–0.49 (–9.85)	0.09 (1.83)

Negative numbers indicate incidents where the opening or closing event (seen in the HSV data) precedes the respective dEGG peak. Opening events were defined as the moment where the initial glottal opening occurred during a glottal cycle. Closing events were defined as the moment where a complete closure was achieved.

axis in both the opening and closing phase (Lohscheller et al., 2013). The EGG signal is a time-varying one-dimensional representation of relative vocal fold contact induced by the complex three-dimensional motion of the vocal folds. Glottal contacting and de-contacting are, strictly speaking, not events that happen at an instant in time (having, theoretically, a duration of 0 s). Rather, they represent phenomena that occur over an interval of time.

The concept of a closing or opening ‘event’ or ‘instant’ thus deserves a more rigorous definition. To avoid confusion, we suggest a distinction between (1) opening/closing instants when considering the glottal airflow and the acoustic excitation in relation to vocal fold vibration; and (2) contacting/de-contacting intervals when describing vocal fold vibratory features and EGG recordings and analyses. When the presumed onset and offset of glottal airflow are being discussed in endoscopic (HSV) data with complete glottal closure, the terms ‘instant of initial glottal opening’ and ‘instant of complete glottal closure’ might be used to indicate those moments in time when the glottal air flow just leaves or reaches the baseline value that is seen during the closed phase (zero in the case of complete glottal closure). With ever-improving technology and increasing HSV frame rates, future research may address the question as to how vertical and A–P phase differences influence the abruptness of glottal airflow cessation during the contacting of the vocal folds, thus influencing the spectral slope of the sound source. In this context it is conceivable that the absence of A–P phase differences (i.e. a more abrupt contacting of the vocal folds) is a prerequisite for optimizing sound generation, e.g. as is needed in un-amplified professional singing (Herbst et al., in press).

The existence of a contacting or de-contacting interval is even more evident if the dEGG signal contains multiple positive or negative peaks. Several previous authors have suggested that these double or multiple dEGG peaks are induced by phase differences along the A–P glottal axis (Hess and Ludwigs, 2000; Henrich et al., 2004; Orlikoff et al., 2012). Analysis of the data gathered in the present study supports these findings: all occurrences of multiple dEGG peaks in either the contacting or de-contacting phase coincided with A–P or hourglass vibratory vocal fold patterns, presumably induced by x-2n or x-3n vibratory modes. However, not all occurrences of hourglass vibratory patterns (see sequence 2) coincided with clear double peaks in the dEGG signal. Further research is necessary to clarify this issue.

In six out of the eight events during glottal opening and closure in this study, a dEGG peak coincided with a dVFCL peak. By its nature, the VFCL is only sensitive to changes in vocal fold contact along the A–P glottal axis, because the video data are acquired from a superior viewpoint. Thus, vertical phase differences (along the inferior–superior dimension) are not reflected in quantitative data gained from the detection of the time-varying glottal edges. In contrast, the EGG signal – measuring the time-varying relative vocal fold contact area – is influenced by phase differences of vocal fold vibration along both the visible A–P and the mostly hidden inferior–superior dimension. Therefore, it is likely that any dEGG peak that coincides with a dVFCL peak is caused by an A–P phase difference. By inversion of this argument, we further hypothesize that any dEGG peak that did not coincide with a dVFCL peak may have been caused by either (1) a vertical phase delay of vocal fold vibration or (2) an inhomogeneity of the vocal fold structure (or thickness) along the glottal axis, e.g. when the vocal processes are involved in the vibration.

Conclusions

In conclusion, the evidence presented in this manuscript does not support the common assumption that the maxima found in the

dEGG signal always coincide with the moments of glottal closure and opening. Contacting and de-contacting of the vocal folds does not occur at an infinitesimally small instant of time, but extends over a certain interval. The duration of the vocal fold contacting and de-contacting intervals are governed by vibratory phase differences along the A–P glottal axis, which have been observed to cause dEGG double peaks. The VFCL was introduced as a promising new parameter for assessing features of vocal fold vibration from HSV data. Further research (employing HSV recordings with maximally achievable frame rates) is needed in order to examine the exact relationship of these contacting and de-contacting intervals to both A–P and inferior–superior phase differences of vocal fold vibration, possibly also analyzing their relationship to glottal airflow and the acoustic output.

MATERIALS AND METHODS

The excised larynx of a female golden retriever (~6 yr and 30 kg body mass), which died of natural causes, was phonated in an excised larynx setup, as described in a previous publication [see supplementary material in Herbst et al. (Herbst et al., 2012)]. The excised larynx was mounted on a vertical air supplying tube. The upper 4 cm of the specimen's trachea formed an airtight seal with that tube. The vocal folds were adducted with three-pronged devices as described in Titze (Titze, 2006). No longitudinal tension was applied on the vocal folds. The larynx was phonated by blowing warmed and humidified air through the adducted glottis. Subglottal air pressure was controlled manually with a Tescom Regulus 3 D50708 pressure valve (McKinney, TX, USA), and was varied between 0 and 20.5 cm H₂O (see Fig. 1), as measured by a Keller PR-41X pressure sensor (Winterthur, Switzerland) positioned 32 cm upstream from the vocal folds.

As regards the appearance of peaks in the dEGG signal, three stereotypical dEGG waveforms can be conceptualized: (A) clear single peaks in both the contacting and the de-contacting phase, (B) a single peak in the de-contacting phase and a pronounced double peak in the contacting phase or (C) a pronounced double peak in the de-contacting phase (and a single peak in the contacting phase). Based on previous research (Hess and Ludwigs, 2000; Henrich et al., 2004), it is hypothesized that scenarios B and C are influenced by A–P phase differences of vocal fold vibration, and scenario A is not.

Following this model, three sequences were selected from a pressure sweep (subglottal pressure ranging from 0 to 20.5 cm H₂O, phonation threshold pressure at ~4 cm H₂O), each representing one of the three stereotypical dEGG waveforms described above. The individual sequences were selected based on visual inspection of the dEGG signal without any prior knowledge of the HSV data, thus precluding human bias in the selection process. Each sequence had to stem from a locally stable region within the signal, having a minimum of 20 similar periods of oscillation.

The EGG signal was captured with a Glottal Enterprises EG 2-1000 two-channel electroglottograph (lower cut-off frequency at 2 Hz, Syracuse, NY, USA). For reference purposes, acoustic recordings were made with a DPA 4061 omni-directional microphone (DPA Microphones, Alleroed, Denmark) positioned 7 cm from the vocal folds. Both the acoustic and the EGG signal were recorded with an RME Fireface 800 external interface (RME, Haimhausen, Germany) at a sampling frequency of 44,100 Hz. The dEGG signal was calculated as the first derivative of the recorded EGG signal, using the formula for the first central difference:

$$\frac{dz}{dt} = \frac{x[i+1] - x[i-1]}{2\Delta t}, \quad (1)$$

where x is the analyzed signal, i is the sample index and Δt is the sampling period.

HSV recordings were made with a Photron FASTCAM 1024 PCI camera (Photron Limited, Tokyo, Japan) at a frame rate of 27,000 images s⁻¹. In order to provide sufficient illumination for such a high frame rate, two light sources were used simultaneously: a dedocool system (Dedo Weigert Film GmbH, Munich, Germany), and a custom built array of twelve 5 W MR16 LED bulbs (SLV Elektronik GmbH, Übach-Palenberg, Germany), powered

by a 12 V car battery. The long-term heat emission from both systems peaked at 31°C, as measured with a Voltcraft IR 260-8S infrared thermometer (Voltcraft, Hirschau, Switzerland).

Synchronization between the HSV and the EGG/acoustic recordings was achieved with a rectangular transistor-transistor logic (TTL) signal (irregular but known pulse duration of ~20 ms, encoding the recording time) generated by a LabJack U6 data acquisition card (LabJack Corporation, Lakewood, CO, USA) that was routed through an IC555 circuit with a rise time of 15 ns. This TTL signal was recorded both as a time-varying voltage by the Fireface sound interface in a separate channel, and as a blinking LED light by the HSV system (see supplementary material Fig. S1). The time-varying intensity values of the pixels in the HSV representing the blinking LED were averaged in each video frame (see supplementary material Fig. S2), and the resulting signal was compared with the TTL signal as captured by the sound card. In cases where the LED took more than one video frame to reach its maximum brightness, the first video frame where the color intensity was greater than the baseline (LED not lit) was chosen to be the onset of the TTL signal. The TTL signal consisted of a steady train of TTL pulses (22 pulses s⁻¹) over the entire duration of the recording, thus ruling out the possibility of a time drift between video system and sound card. The TTL synch signals for both the EGG and the HSV data were correlated to each other by a supervised semi-automatic procedure that is outlined in supplementary material Fig. S3. The synchronization accuracy was dependent on the video frame rate (i.e. the lower of the two sampling frequencies involved). The maximum synchronization error was calculated to be 0.037 ms, i.e. the time delay between two consecutive video frames at a video frame rate of 27,000 frames s⁻¹.

Analysis

In digital kymography (DKG) (Wittenberg et al., 2000), the principles of videokymography are applied to HSV sequences. In order to create a DKG, a line perpendicular to the vocal fold axis is selected within a HSV sequence, and the corresponding video pixels on that line are successively extracted for each video frame in the analyzed sequence. The extracted lines are concatenated in time (separated by the frame rate period) to form the final graph (Švec and Schutte, 2012). The DKGs created for this manuscript were generated using a custom-written Python script (Herbst, 2012), which was run as a plug-in within the FIJI image analysis software package (Schindelin et al., 2012). DKGs were extracted at four equidistantly spaced positions along the glottal axis (Orlikoff et al., 2012) in order to visualize the different vibratory patterns along the glottal axis (Orlikoff et al., 2012; Lohscheller et al., 2013).

To enable quantitative analysis of the vibrating patterns along the entire length of the vocal folds, a clinically evaluated image processing procedure was applied, which is described in detail elsewhere (Lohscheller et al., 2007). With this algorithm, the medial edges of both vocal folds were extracted within each frame of the HSV recording. The segmentation results were superimposed upon the DKGs shown in Figs 2, 4 and 5.

A new parameter, the VFCL, was defined as a measure of the relative degree of vocal fold contact along the A–P glottal axis (Herbst et al., 2013). The VFCL was calculated based on the previously extracted time-varying glottal edge data. This parameter is sensitive to A–P phase differences of vocal fold vibration, but not to vertical phase differences. The VFCL parameter is in essence similar to the membranous contact quotient (MCQ) introduced by Scherer et al. (Scherer et al., 1997). The VFCL differs from the MCQ in that it is calculated for every frame in the HSV data (i.e. the VFCL is a dynamic parameter), whereas the MCQ is determined on a cycle-to-cycle basis, considering the maximum closure along the A–P glottal axis.

The extracted time-varying glottal edges were further used to create GVGs (see Karakozoglou et al., 2012), a visualization technique that transfers information on the time-varying glottal width (as color information) along the A–P dimension into a single graph (Lohscheller et al., 2008). In a GVG, time is displayed on the x-axis, the A–P glottal axis is shown on the y-axis, and the respective normalized distance of the left and right glottal edge (in pixels) is depicted as color information on the z-axis. The GVG can be used to objectively describe the two-dimensional vibration type of glottal opening and closure. For creating the GVG plots shown in this manuscript, no interpolation was used by the plotting software to map

the GVG data to the individual pixel values in the graphs. To increase the visibility of smaller vocal fold edge distances within the generated GVGs, the normalized GVG z-axis values were transformed to logarithmic values using the formula $\zeta_{[x,y]} = \log_{10}(9z_{[x,y]} + 1)$.

Acknowledgements

We kindly thank R. Hofer for contributing to the setup of the excised larynx experiment. We are very thankful to the reviewers for their time and expertise, and for their insightful comments.

Competing interests

The authors declare no competing financial interests.

Author contributions

C.T.H. conceived and wrote the manuscript, and led the data analysis. C.T.H. and W.T.F. created the setup for the excised larynx experiment. G.W. provided and partly prepared the canine larynx specimen. C.T.H. and W.T.F. conducted the excised larynx experiment. J.L. contributed to the analysis of the high-speed video data. J.L. and J.G.Š. contributed significantly to the interpretation of the collected data. J.G.Š., J.L., N.H. and W.T.F. contributed significantly to critically revising the manuscript.

Funding

This research was supported by European Research Council Advanced Grant 'SOMACCA' and a start-up grant from the University of Vienna (to C.T.H. and W.T.F.); the European Social Fund and the state budget of the Czech Republic, project nos CZ.1.07/2.3.00/30.0004 'POST-UP' (to C.T.H. and J.G.Š.) and OPVK CZ.1.07/2.3.00/20.0057 (to J.G.Š.); and grant no. LO1413/2-2 by the Deutsche Forschungsgemeinschaft (to J.L.).

Supplementary material

Supplementary material available online at <http://jeb.biologists.org/lookup/suppl/doi:10.1242/jeb.093203/-/DC1>

References

- Anastaplo, S. and Karnell, M. P. (1988). Synchronized videostroboscopic and electroglottographic examination of glottal opening. *J. Acoust. Soc. Am.* **83**, 1883–1890.
- Baer, T. (1981). Observation of vocal fold vibration: measurements of excised larynges. In *Vocal Fold Physiology* (ed. K. N. Stevens and M. Hirano), pp. 119–133. Tokyo: University of Tokyo Press.
- Baer, T., Löfqvist, A. and McGarr, N. S. (1983). Laryngeal vibrations: a comparison between high-speed filming and glottographic techniques. *J. Acoust. Soc. Am.* **73**, 1304–1308.
- Baken, R. J. (1992). Electroglottography. *J. Voice* **6**, 98–110.
- Baken, R. J. and Orlikoff, R. F. (2000). *Clinical Measurement of Speech and Voice*, 2nd edn. Toronto, ON: Singular Publishing, Thompson Learning.
- Berke, G. S., Moore, D. M., Hantke, D. R., Hanson, D. G., Gerratt, B. R. and Burstein, F. (1987). Laryngeal modeling: theoretical, *in vitro*, *in vivo*. *Laryngoscope* **97**, 871–881.
- Berry, D. A., Herzel, H., Titze, I. R. and Krischer, K. (1994). Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions. *J. Acoust. Soc. Am.* **95**, 3595–3604.
- Bless, D. M., Hirano, M. and Feder, R. J. (1987). Videostroboscopic evaluation of the larynx. *Ear Nose Throat J.* **66**, 289–296.
- Childers, D. G. and Krishnamurthy, A. K. (1985). A critical review of electroglottography. *Crit. Rev. Biomed. Eng.* **12**, 131–161.
- Childers, D. G., Naik, J. M., Larar, J. N., Krishnamurthy, A. K. and Moore, G. P. (1983). Electroglottography, speech, and ultra-high speed cinematography. In *Vocal Fold Physiology and Biophysics of Voice* (ed. I. R. Titze and R. Scherer), pp. 202–220. Denver, CO: Denver Center of Performing Arts.
- Childers, D. G., Hicks, D. M., Moore, G. P. and Alsaka, Y. A. (1986). A model for vocal fold vibratory motion, contact area, and the electroglottogram. *J. Acoust. Soc. Am.* **80**, 1309–1320.
- Deliyski, D. D. and Hillman, R. E. (2010). State of the art laryngeal imaging: research and clinical implications. *Curr. Opin. Otolaryngol. Head Neck Surg.* **18**, 147–152.
- Deliyski, D. D., Petrushev, P. P., Bonilha, H. S., Gerlach, T. T., Martin-Harris, B. and Hillman, R. E. (2008). Clinical implementation of laryngeal high-speed videendoscopy: challenges and evolution. *Folia Phoniatr. Logop.* **60**, 33–44.
- Fabre, P. (1957). Un procédé électrique percutané d'inscription de l'accollement glottique au cours de la phonation: glottographie de haute fréquence; premiers résultats (A non-invasive electric method for measuring glottal closure during phonation: high frequency glottography; first results). *Bull. Acad. Natl. Med.* **141**, 66–69.
- Flanagan, J. and Landgraf, L. L. (1968). Self oscillating source for vocal tract synthesizers. *IEEE Trans. Audio Electroacoust.* **16**, 57–64.
- Fourcin, A. J. and Abberton, E. (1971). First applications of a new laryngograph. *Med. Biol. Illus.* **21**, 172–182.

- Golla, M. E., Deliyski, D. D., Orlikoff, R. F. and Moukalled, H. (2009). Objective comparison of the electroglottogram to synchronous high-speed images of vocal-fold contact during vibration. In *Proceedings of the 6th International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications MAVEBA*, Vol. 6 (ed. C. Manfredi), pp. 141-144. Firenze, Italy: Firenze University Press.
- Henrich, N., d'Alessandro, C., Doval, B. and Castellengo, M. (2004). On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *J. Acoust. Soc. Am.* **115**, 1321-1332.
- Henrich, N., d'Alessandro, C., Doval, B. and Castellengo, M. (2005). Glottal open quotient in singing: measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency. *J. Acoust. Soc. Am.* **117**, 1417-1430.
- Herbst, C. T. (2012). DKG plugin for FIJI. Available at <http://homepage.univie.ac.at/christian.herbst/index.php?page=fiji>.
- Herbst, C. and Ternström, S. (2006). A comparison of different methods to measure the EGG contact quotient. *Logoped. Phoniater. Vocol.* **31**, 126-138.
- Herbst, C. T., Fitch, W. T. and Švec, J. G. (2010). Electroglottographic wavegrams: a technique for visualizing vocal fold dynamics noninvasively. *J. Acoust. Soc. Am.* **128**, 3070-3078.
- Herbst, C. T., Stoeger, A. S., Frey, R., Lohscheller, J., Titze, I. R., Gumpenberger, M. and Fitch, W. T. (2012). How low can you go? Physical production mechanism of elephant infrasonic vocalizations. *Science* **337**, 595-599.
- Herbst, C. T., Fitch, W. T., Lohscheller, J. and Švec, J. G. (2013). Estimation of the vertical glottal shape based on empirical high-speed video and electroglottographic data. In *Proceedings of the 10th International Conference on Advances in Quantitative Laryngology, Voice and Speech Research* (ed. D. D. Deliyski), pp. 75-76. Cincinnati, OH: Communication Sciences Research Center, Cincinnati Children's Hospital Medical Center.
- Herbst, C. T., Howard, D. M. and Švec, J. G. (in press). The sound source in singing – basic principles and muscular adjustments for fine-tuning vocal timbre. In *The Oxford Handbook of Singing* (ed. G. Welch, D. M. Howard and J. Nix). Oxford, UK: Oxford University Press.
- Hertegård, S. (2005). What have we learned about laryngeal physiology from high-speed digital videendoscopy? *Curr. Opin. Otolaryngol. Head Neck Surg.* **13**, 152-156.
- Hess, M. M. and Ludwigs, M. (2000). Strobophotoglottographic transillumination as a method for the analysis of vocal fold vibration patterns. *J. Voice* **14**, 255-271.
- Higgins, M. B. and Schulte, L. (2002). Gender differences in vocal fold contact computed from electroglottographic signals: the influence of measurement criteria. *J. Acoust. Soc. Am.* **111**, 1865-1871.
- Howard, D. M. (1995). Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers. *J. Voice* **9**, 163-172.
- Kania, R. E., Hans, S., Hartl, D. M., Clement, P., Crevier-Buchman, L. and Brasnu, D. F. (2004). Variability of electroglottographic glottal closed quotients: necessity of standardization to obtain normative values. *Arch. Otolaryngol. Head Neck Surg.* **130**, 349-352.
- Karakozoglou, S.-Z., Henrich, N., d'Alessandro, C. and Stylianou, Y. (2012). Automatic glottal segmentation using local-based active contours and application to glottovibrometry. *Speech Commun.* **54**, 641-654.
- Krenmayr, A., Wöllner, T., Supper, N. and Zorowka, P. (2012). Visualizing phase relations of the vocal folds by means of high-speed videendoscopy. *J. Voice* **26**, 471-479.
- La, P. and Sundberg, J. (2012). Effect of subglottal pressure variation on the "closed quotient" – comparing data derived from electroglottograms and from flow glottograms. In *The Voice Foundation's 41st Annual Symposium: Care of the Professional Voice* (ed. M. Russo), pp. 172. Philadelphia, PA: The Voice Foundation.
- Lohscheller, J., Toy, H., Rosanowski, F., Eysholdt, U. and Döllinger, M. (2007). Clinically evaluated procedure for the reconstruction of vocal fold vibrations from endoscopic digital high-speed videos. *Med. Image Anal.* **11**, 400-413.
- Lohscheller, J., Eysholdt, U., Toy, H. and Dollinger, M. (2008). Phonovibrometry: mapping high-speed movies of vocal fold vibrations into 2-D diagrams for visualizing and analyzing the underlying laryngeal dynamics. *IEEE Trans. Med. Imaging* **27**, 300-309.
- Lohscheller, J., Švec, J. G. and Döllinger, M. (2013). Vocal fold vibration amplitude, open quotient, speed quotient and their variability along glottal length: kymographic data from normal subjects. *Log. Phon. Vocol.* **38**, 182-192.
- McClellan, J. H., Schafer, R. W. and Yoder, M. A. (1998). *DSP First: a Multimedia Approach*. Upper Saddle River, NJ: Prentice-Hall.
- Moore, G. P., White, F. D. and Von Leden, H. (1962). Ultra high speed photography in laryngeal physiology. *J. Speech Hear. Disord.* **27**, 165-171.
- Orlikoff, R. F. (1991). Assessment of the dynamics of vocal fold contact from the electroglottogram: data from normal male subjects. *J. Speech Hear. Res.* **34**, 1066-1072.
- Orlikoff, R. F., Golla, M. E. and Deliyski, D. D. (2012). Analysis of longitudinal phase differences in vocal-fold vibration using synchronous high-speed videendoscopy and electroglottography. *J. Voice* **26**, 816.e13-20.
- Roads, D. and Strawn, J. (1998). Digital Audio Concepts. In *The Computer Music Tutorial* (ed. C. Roads), pp. 5-47. Cambridge, MA: The MIT Press.
- Rothenberg, M. (1979). Some relations between glottal air flow and vocal fold contact area. In *Proceedings of the Conference on the Assessment of Vocal Pathology*, Vol. ASHA Reports No. 11 (ed. C. L. Ludlow and M. O. Hart), pp. 88-96. Rockville, MD: American Speech and Hearing Association.
- Rothenberg, M. and Mahshie, J. J. (1988). Monitoring vocal fold abduction through vocal fold contact area. *J. Speech Hear. Res.* **31**, 338-351.
- Rubin, H. J. and Le Cover, M. (1960). Technique of high-speed photography of the larynx. *Ann. Otol. Rhinol. Laryngol.* **69**, 1072-1082.
- Sapienza, C. M., Stathopoulos, E. T. and Dromey, C. (1998). Approximations of open quotient and speed quotient from glottal airflow and EGG waveforms: effects of measurement criteria and sound pressure level. *J. Voice* **12**, 31-43.
- Scherer, R. C., Druker, D. G. and Titze, I. R. (1988). Electroglottography and direct measurement of vocal fold contact area. In *Vocal Fold Physiology: Voice Production, Mechanisms and Functions*, Vol. 2 (ed. O. Fujimura), pp. 279-290. New York, NY: Raven Press.
- Scherer, R. C., Alipour, F., Finnegan, E. and Guo, C. G. (1997). The membranous contact quotient: a new phonatory measure of glottal competence. *J. Voice* **11**, 277-284.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B. et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676-682.
- Schutte, H. K. and Miller, D. G. (2001). Measurement of closed quotient in a female singing voice by electroglottography and videokymography. In *Vth International Conference Advances in Quantitative Laryngology, Groningen, the Netherlands, April 27-28, 2001*. CD-ROM (ed. H. K. Schutte). Groningen the Netherlands: Groningen Voice Research Laboratory, University of Groningen.
- Švec, J. G. (2000). *On Vibration Properties of Human Vocal Folds: Voice Registers, Bifurcations, Resonance Characteristics, Development and Application of Videokymography*. Doctoral dissertation. Groningen, the Netherlands: University of Groningen.
- Švec, J. G. and Schutte, H. K. (1996). Videokymography: high-speed line scanning of vocal fold vibration. *J. Voice* **10**, 201-205.
- Švec, J. G. and Schutte, H. K. (2012). Kymographic imaging of laryngeal vibrations. *Curr. Opin. Otolaryngol. Head Neck Surg.* **20**, 458-465.
- Švec, J. G., Sundberg, J. and Hertegård, S. (2008). Three registers in an untrained female singer analyzed by videokymography, strobolaryngoscopy and sound spectrography. *J. Acoust. Soc. Am.* **123**, 347-353.
- Tanabe, M., Kitajima, K., Gould, W. J. and Lambiase, A. (1975). Analysis of high-speed motion pictures of the vocal folds. *Folia Phoniatr. (Basel)* **27**, 77-87.
- Teaney, D. and Fourcin, A. (1980). The electrolaryngograph as a clinical tool for the observation and analysis of vocal fold vibration. In *Ninth Symposium Care of the Professional Voice*, pp. 128-134. New York, NY: The Juilliard School, The Voice Foundation.
- Titze, I. R. (1989). A four-parameter model of the glottis and vocal fold contact area. *Speech Commun.* **8**, 191-201.
- Titze, I. R. (1990). Interpretation of the electroglottographic signal. *J. Voice* **4**, 1-9.
- Titze, I. R. (2006). *The Myoelastic Aerodynamic Theory of Phonation*. Denver, CO: National Center for Voice and Speech.
- Titze, I. R., Jiang, J. J. and Hsiao, T. Y. (1993). Measurement of mucosal wave propagation and vertical phase difference in vocal fold vibration. *Ann. Otol. Rhinol. Laryngol.* **102**, 58-63.
- Wittenberg, T., Tigges, M., Mergell, P. and Eysholdt, U. (2000). Functional imaging of vocal fold vibration: digital multislice high-speed kymography. *J. Voice* **14**, 422-442.
- Yamauchi, A., Imagawa, H., Sakakibara, K., Yokonishi, H., Nito, T., Yamasoba, T. and Tayama, N. (2013). Phase difference of vocally healthy subjects in high-speed digital imaging analyzed with laryngotopography. *J. Voice* **27**, 39-45.